Decoupled Execution Paradigm for Data-Intensive High-End Computing

**Motivating Example**
- Data commonly represented by a multi-dimensional array-based data model
- Read required data from storage servers to compute nodes
- Perform computations on desired data with specified conditions, and then write back
- With clear data retrieval and processing phases and computing and simulation phases
- Data access and movement often dominate execution time for data-intensive HEC apps

**Programming Model**
- To determine operations to be passed to data nodes
- Designed as an MPI extension, allowing users to specify operations conducted on data nodes
- Results sent back to compute nodes for further processing

**Runtime System**
- Relies on two libraries, message passing library and data processing library
- Message passing library focuses on the memory abstraction and provides support for computation-intensive operations
- Leverage the existing MPI library for this purpose
- Data processing library focuses on the I/O abstraction and provides support for data-intensive operations

**Experimental Results**
- Two application kernels
  - Kernel calculation of the CESM that computes the moving average of selected area of specified data
  - Flow routing and flow accumulation calculations in geographic information system

**Execution Paradigm Comparison**
- Conventional Execution Paradigm
- Decoupled Execution Paradigm

**I/O Bottleneck**
Significant gap between computing and I/O.

Long I/O latency leads to performance degradation.

**System Architecture**
Decoupled Execution Programming Model (DEPM) (MPI Extension)

- **Applications**
- **Decoupled Execution Systems Architecture**
- **Compute Nodes**
- **Local SSD storage**
- **Data Nodes**
- **Storage-side Data Nodes**
- **Compute-side Data Nodes**

**Execution Paradigm**
- **Bottleneck**
- **Retrieval**
- **Reduction**
- **Compute**
- **Store**

**Acknowledgment**
This research is sponsored in part by the U.S. National Science Foundation under grants CNS-1162540, CNS-1162488, CNS-1161507 and by the NSF grant CCF-0937877.